Graphics, Vision and Image Processing Journal, ISSN 1687-398X, Volume 16, Issue 2, ICGST LLC, Delaware, USA, Dec. 2016

www.icgst.com

Efficient Color Image Segmentation relying on Binary Image Interest Points

N. El aryby[1], W. Aitfares[1,2], F. Regragui[1]
[1]*LIMIARF- FSR, University Mohammed V- Rabat B.P 1014 RP, Rabat, Morocco.*
[2] *GENIUS laboratory, SUPMTI, 98, Avenue Allal Ben Abdellah, Hassan-Rabat, Morocco.*
elaryby@gmail.com, aitfares@supmti.ma, regragui@fsr.ac.ma

Abstract
The present work proposes a new method for color image segmentation. This approach is based on the calculation of the interest points from the binary image to extract the binary mask used to separate the object from its background. The binarization of the image is done by using the local Otsu threshold method to reduce the effect of heterogeneity thus improving the object segmentation. To evaluate the effectiveness of the proposed method, we compared the obtained results of the object segmentation based on our approach with the results of the grab cut method. Experimental results show the robustness of our approach for detecting objects in real images.

Keywords: *Color image segmentation, Lab color space, heterogeneity, local Otsu thresholding, interest points, morphological processing.*

1. Introduction
This paper addresses the problem of object segmentation which is a challenging task especially in the case of the presence of heterogeneity in the image. Segmentation techniques can be classified into different categories [1]: Threshold based, Region-based, Cluster-based and Edge-based.
Image segmentation based on the thresholding is one of the oldest techniques. It classifies the pixels into two classes: pixels which have intensity value less than a threshold belong to one class; while other pixels belong to the other class [2, 3]. Region-based methods classify an image into different regions where pixels within each region are sharing similar properties according to predefined conditions [4]. Clustering is an unsupervised learning technique where the number of clusters is determined in advance to classify pixels and those having similar conditions are grouped together into one cluster [5]. Edge-based segmentation methods attempt to segment the image by detecting the edges between different regions [6]. These edges are detected in locations where a significant intensity change occurs. The pixels in these locations are extracted and grouped to highlight the edges in the image. Object segmentation can also be accomplished by the active contour method which has become very popular and widely used in image segmentation. This method has two approaches, edge-based and region-based. In the edge-based active contour approach [7, 8], image gradients are used to evolve a contour until it reaches the object boundaries; nevertheless such a method suffers from sensitivity to noise and active contour initialization. Alternatively, the region-based active contour approach [9, 10, 11, 12, 13, 14] considers the information extracted from the region instead of the image gradients. These approaches model the foreground and the background regions statistically and deform the active contour to identify the real object boundaries. For both edge-based and region-based active contour, an energy function is minimized to allow the contour to deform and segment the object of interest when the optimum of this energy is reached. However, these approaches still suffer from limitations due essentially to parameter estimation and the position of the initial active contour in the image.
Heterogeneity is almost omnipresent in real images on both the object and the background which hinders the process of the object segmentation and makes it more difficult. To reduce difficulties due to heterogeneity, we propose in this paper a new segmentation method relying on interest points extracted from the binary image rather than the real image where a local thresholding is used. Interest points are particularly relevant because they are simple and robust low-level features providing an efficient characterization of the object. They can be defined as being points in the image where significant changes occur, such as corners, connections, black points on a white background or other points marked by an important change of texture. In our work, we use the Harris detector [15] for its popularity and its satisfactory results in extracting the interest points.
The paper is organized as follows. Thresholding is defined in section 2. The Harris detector is described in section 3. Morphological operations are defined in section 4. We present the proposed object segmentation

approach in section 5 and experimental results in section 6. We conclude the paper in section 7.

2. Thresholding

The purpose of thresholding methods is the reduction of unnecessary information in the image leaving only the useful one for further processing. It is the process of classifying the image into two regions corresponding to the object and its background using an optimum threshold value. We categorize the thresholding methods into two groups: global and local methods. The global methods use a single global threshold value to segment an image into two classes such as the one proposed by Otsu [16] and Kapur and al. [17], whereas local methods estimate a threshold value for each pixel according to the information extracted from the neighboring pixels such as range, variance or surface-fitting parameters. The techniques of Bernsen [18], Chow and Kaneko [19], Eikvil [20], Mardia and Hainsworth [21], Niblack [22], Taxt [23], Yanowitz and Bruckstein [24] and Sauvola and Pietikainen [25] belong to this category.

a. Global thresholding

Otsu's thresholding is a clustering method which selects the optimal threshold by maximizing the inter-class variance (between class), which is same as minimizing the intra-class variance (within class). The within class variance can be formulated as:

$$\sigma^2_{within}(T) = n_B(T)\sigma^2_B(T) + n_O(T)\sigma^2_O(T). \quad (1)$$

where T is the threshold value. $n_B(T)$ and $n_O(T)$ denote respectively the sum of the pixel intensities of the background and the object. $\sigma^2_B(T)$ and $\sigma^2_O(T)$ are respectively the variance of the pixels in the background and the object.
The between class variance is given as:

$$\sigma^2_{between}(T) = n_B(T)n_O(T)\left[\mu_B(T) - \mu_O(T)\right]^2. \quad (2)$$

The optimal threshold can be defined as:

$$T_{opt} = \arg\max(\sigma^2_{between}(T)) = \arg\min(\sigma^2_{within}(T)). \quad (3)$$

where μ_B and μ_O represent respectively the mean value of the pixels intensities in the background and the object region.
Otsu's method gives satisfactory results when the region of each class is homogeneous in terms of pixel intensities; this method still remains one of the most popular thresholding methods. Based on the entropy, Kapur's method perform bi-level image thresholding. in which the foreground and background of the image are considered as two different signal sources. In this way, ~~so that~~ when the sum of the two class entropies reaches its maximum, the image is considered to be optimally thresholded.. Let I be an image containing n pixels described by gray levels belonging to the set {0,1,… L-1} with probability distribution $p_i = p_1, p_2, ..., p_L$. From this distribution, we derive two probability distributions, one for the object (f) and the one for the background (b), given by:

$$p_f = \sum_{i=1}^{t} p_i \quad , \quad p_b = \sum_{i=t+1}^{L} p_i \quad (4)$$

Where t is the threshold
Using the expressions of the foreground and background entropies defined respectively by:

$$H_f = -\sum_{i=0}^{t} \frac{p_i}{p_f} \log \frac{p_i}{p_f}. \quad (5)$$

$$H_b = -\sum_{i=t+1}^{L-1} \frac{p_i}{p_b} \log \frac{p_i}{p_b}. \quad (6)$$

The optimal threshold is the value maximizing the aggregated entropy:

$$T_{opt} = \arg\max[H_f + H_b]. \quad (7)$$

b. Local adaptive thresholding

In the method from Sauvola [26] which is an improvement of the Niblack's [22] method, the threshold $T(x,y)$ is calculated using the mean $m(x,y)$ and the standard deviation $\delta(x,y)$ of the pixel intensities within a window of size $w \times w$ as:

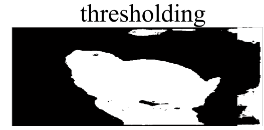
$$T(x,y) = m(x,y)\left[1 + k(\frac{\delta(x,y)}{R} - 1)\right]. \quad (8)$$

R is the maximum value of the standard deviation and k is a bias term which takes positive values in the range [0.2, 0.5].
In general, each global method can be applied in a local version by dividing the image into blocks and then using a global threshold on each block independently. Figure1 shows an example of image thresholding using various methods of binarization previously defined.


(a) Input image


(b) global Otsu thresholding


(c) global entropy thresholding


(d) local otsu thresholding(180x180 pixels)


(e) local entropy thresholding (180x180 pixels)

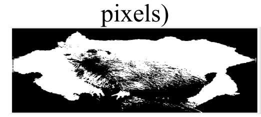

(f) Sauvola Thresholding(180x180 pixels)

Figure 1. Thresholding methods.

From the results presented in Figure1, it can be inferred that, the local Otsu's method provides good thresholding compared to other local and global methods.

3. Harris Detector

Extracting the interest points by the Harris detector is a method based on the Moravec detector [26]. The method extracts the corners as interest points by using a differential method. Harris detector is based on the autocorrelation of the image intensity values or the image gradient values. The gradient covariance matrix is given by:

$$C = \begin{bmatrix} (\frac{\partial I}{\partial x})^2 & \frac{\partial I}{\partial x}\frac{\partial I}{\partial y} \\ \frac{\partial I}{\partial y}\frac{\partial I}{\partial x} & (\frac{\partial I}{\partial y})^2 \end{bmatrix} = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}. \quad (9)$$

where I_x and I_y denote the image gradients in the x and y directions. The Harris detector considers the minimum and the maximum eigenvalues, respectively α and β, of the image gradient covariance matrix C in developing the corner detector. A 'corner' occurs when the two eigenvalues are large and similar in magnitude. Harris [27] proposes a measure using the determinant and the trace of the gradient covariance matrix defined as:

$$H = \alpha\beta - k(\alpha+\beta)^2 = \det C - k(\mathrm{Trace}(C))^2. \quad (10)$$

where k belongs to the interval [0,04 0,06].

The pixels are classified according to the values of H such as:

$H > 0$: Corner pixel, $H \sim 0$: pixel in a flat region and
$H < 0$: edge pixel.

4. Morphological operations

Morphology operations [28] are techniques of the image processing which depend on the shape. The value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. The fundamental morphological operations are erosion and dilation. Erosion is typically applied to binary images, but they are versions that work on grayscale images. The basic effect of the operator on a binary image is to erode away the boundaries of the image foreground. The value of the output pixel is the minimum value of all the neighboring pixels in the input image. The erosion operation is defined by the function:

$$f(x,y)=\min\{I(x,y) \text{ and its neighboring pixels}\}. \quad (11)$$

$f(x,y)$ is the eroded image at pixel (x,y), and I(x,y) is the pixel intensity.

Dilation adds pixels to the boundary of the regions of foreground pixels. Thus, areas of foreground regions grow in size while holes within those regions become smaller. The value of the output pixel is the maximum of all the neighboring pixels in the input image. The dilation is defined by:

$$g(x,y)= \max\{I(x,y) \text{ and its neighboring pixels}\}. \quad (12)$$

$g(x,y)$ is the dilated image at pixel (x,y).

5. Proposed Approach

In this section, we describe the proposed object segmentation approach. We use the interest points extracted from the binary image after a thresholding step as demonstrated in the flowchart presented in Figure 2. The region with the maximum number of interest points is selected to be the zone containing the object of interest. By calculating the binary mask, the object is well separated from its background and the object segmentation is achieved. The contour of the object is then detected on the input image.

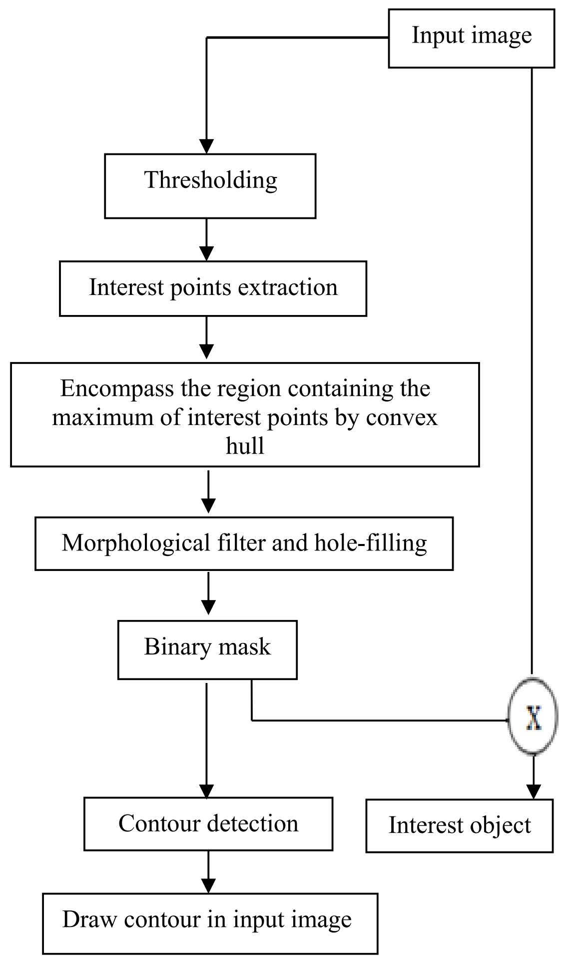


Figure 2. Flowchart of the proposed algorithm.

To test the efficiency of this algorithm on real images, let consider the image of the boat presented in Figure 3(a). The binarization of the image in our work (Figure 3(b)) is obtained by using the local Otsu's thresholding technique (180 x 180 pixels) which produces satisfactory results compared to the global thresholding methods as well as the local methods as shown in the Figure 1. In the local Otsu's method, the threshold is computed individually for each pixel using global thresholding from the local neighborhood of the pixel. This operation classifies the image into regions (clusters) separating the

foreground image from the background image. The interest points are extracted (Figure 3(c)) from the binary image. These points are local maxima and differ from their neighbors in intensity, color and curvature direction. In addition, they are originally utilized to characterize the areas with the most visual information. The interest object is designed to be the most dominant region in terms of information and thus to be in the cluster containing the maximum number of interest points. In our work, a convex hull (red curve) is drawn to delimit this interest region as shown in Figure 3(d). Morphological operations with a filter are applied on the binary image. Dilation is used to enlarge the boundaries of the foreground region while filtering is used to fill the holes in the interest region (Figure 3(e)). The binary mask is then obtained as shown in Figure. 3(f). The latter is multiplied by the input real image to separate the object from its background (Figure. 3(g)). The boundary of the object is finally detected and drawn in the image as shown in Figure 3(h). The energy of our segmentation method can be expressed as:

$$E = -\left(\arg\max(\sigma^2_{between}(T)) + \max(H) + g(x,y) \right). \qquad (13)$$

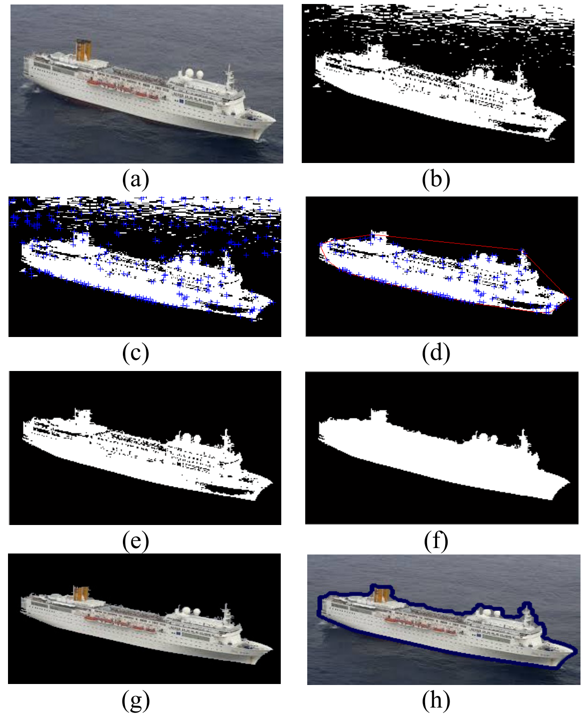


Figure 3. Image segmentation steps.

6. Experiments

(a) (b) (c) (d)

Figure 4. (a) Input image, (b) Binary mask, (c) Segmented interest object, (d) Object of interest boundary (blue curve).

To evaluate the performance of the proposed approach, we used various pictures from the Berkeley image database [29]. In our work, we consider the RGB image, as an input image. The RGB image is converted to Lab color space. We adopt the CIE L*a*b* color feature here since this color space is one of the most widely adopted color models for describing colors visible to humans. Figure 4 shows the results obtained by the proposed approach. The column (a) of this figure represents the real input images with heterogeneous characteristics. The column (b) represents the binary mask obtained after the morphological filter and hole-filling operations. On one hand, the binarization method preserves the maximum information in the image; on the other hand, the interest points allow an efficient characterization of the interest region in the binary image. As a result, an efficient object segmentation can be obtained as shown in the images displayed in column (c). In each case, the obtained result shows the efficiency of our proposed approach in terms of accuracy of the object segmentation even if the heterogeneity appears in the foreground image (second and seventh rows of figure 4) or in the background image (first and fourth rows of the figure 4) or on both of them to reveal some advantages of our proposed approach compared to the active contour method. It is found to be robust in finding directly the real object boundary (column (d) of Figure 4) without any evolution of a curve (active contour) as required with the active contour method. In fact with this method, an automatic initialization of the active contour can be provided by the curve presented by the convex hull around the interest object (Figure.3 (d)). Then, the evolution of the active contour is made according to an iterative process of deformation of this curve until this latter reaches the object boundary.

We have compared the proposed approach with GrabCut method [30] which separates object from the background in color image under some constraints. The segmentation process is realized by defining manually an initial rectangle around the object of interest. All the pixels belonging to the outside of the rectangle are assumed to belong to the background region, and those belonging to the inside of the rectangle are assumed to define the object region (with a part belonging to the background). With an iterative method, the image pixels are finally classified into two clusters defining either the object region or the background region.

Input image GrabCut proposed method

Figure 5. Comparison of the segmentation results obtained by the GrabCut and the proposed method.

All the algorithms used were implemented in MATLAB and tested on a Intel (R) Core (TM) i5-2520M 2.50GHz CPU, 4GB Memory, Windows 7.

It is clear from the results presented in Figure5 that our proposed approach shows better performance than the grab cut method. Indeed in the five images of figure5 named, respectively, “rower”, “tiger”, “horses”, “mollusk” and ”echinoderm”, the object of interest is separated accurately from the background image using our approach. The GrabCut method with an initialization around the interest object produces some falsely detected background pixels assigned to the foreground (first, third and last rows of Figure 5).

We also compared the energy convergence in term of the execution time for both the proposed approach and the GrabCut method. The results in Figure 6 concern the echinoderm image (the third row of Figure 5) as an example of image presenting heterogeneous characteristics in both foreground and background image. The energy calculated with our approach (red curve) converges faster (in 0.4 seconds) than that calculated with GrabCut method (blue curve) requiring more time to converge (0.8 seconds). (Figure 5) and also in term of computation time (Figure 6).

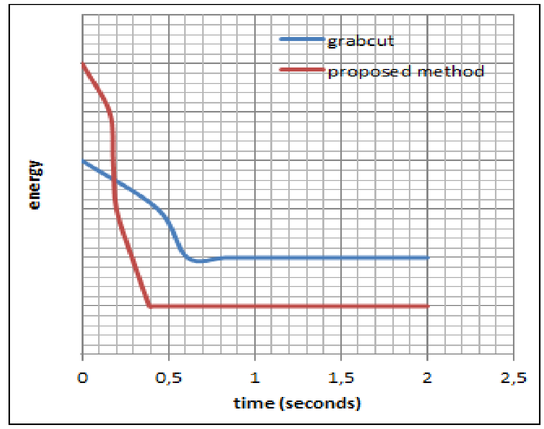


Figure 6. Execution time of echinoderm image segmentation using GrabCut and proposed method.

A statistical comparison of the segmentation results obtained by the proposed method and the GrabCut method relying on similarity PSNR (Peak Signal-to-Noise Ratio) and MSE (Mean Square Error) is presented in tables 1 and 2. From these tables, our method shows that the PSNR measure maintains a high value for all images compared to the PSNR measured for the GrabCut method. Similarly, lower value of the error (MSE) was obtained with our approach compared to the one obtained with the grab cut method.
Consequently, our proposed approach is revealed more efficient than the GrabCut method; it provides more accuracy in performing correct object segmentation.

Image	Metrics (db)	Proposed method	GarbCut method
rower	PSNR	26,61	26,15
	MSE	0,59	0,60
tiger	PSNR	28,40	25,35
	MSE	0,13	0,78
horses	PSNR	26,47	25,51
	MSE	0,60	0,75
mollusk	PSNR	25,53	25,43
	MSE	0,76	0,77
echinoderm	PSNR	25,41	25,12
	MSE	0,78	0,83

Table 1. Performance comparison of GrabCut and the proposed method using PSNR and MSE.

Metrics(db)	Proposed method	GrabCut method
PSNR	26,48	25,51
MSE	0,57	0,75

Table 2.Average performance of GrabCut and the proposed method using PSNR and MSE.

7. Conclusion

In this paper, a new approach is proposed for color object segmentation. The object is extracted from the background image relying on binary image interest points, and then the object boundary is detected. The proposed method is implemented using local thresholding, interest points and morphological operations. Experimental results have proven its competitiveness with well-known

methods in literature especially with the active contour and grab cut methods. In future work, we will extend our work to detect and track multiple objects.

References

[2] D. Rajeshwar, Priyanka, D. Swapna. Image Segmentation Techniques. International Journal on Electronics & Communication Technology (IJECT), Vol. 3,Issue 1, (2012) 66-70

[3] W.X. Kang, Q.Q. Yang, R.R. Liang. The Comparative Research on Image Segmentation Algorithms. IEEE Conference on ETCS (2009) 703-707.

[4] Md. Habibur Rahman, Md. Rafiqul Islam, “Segmentation of Color Image using Adaptive Thresholding and Masking with Watershed Algorithm”, IEEE, 2013.

[5] H.G. Kaganami, Z. Beij. Region Based Detection versus Edge Detection. IEEE Transactions on Intelligent information hiding and multimedia signal processing (2009) 1217-1221.

[6] V.K. Dehariya, S.K. Shrivastava, R.C. Jain. Clustring of Image Data Set Using K-means and Fuzzy K-Means Algorithms. International conference on CICN (2010) 386-391

[7] E. Argyle. Techniques for Edge Detection. Proc. IEEE, Vol. 59, (1971) 285-286

[8] V. Caselles, R. Kimmel, G. Sapiro. Geodesic Active Contours. Int. J. Comput. Vision, Vol. 22,Issue 1 (1997) 61-79

[9] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active Contour Models,” Int. J. Comput. Vision, vol. 1, Issue. 4, (1988) 321-332.

[10] T.Chan, L. Vese. Active Contours Without Edges. IEEE Trans.Image Process,Vol. 10,Issue 2 (2001) 266-277

[11] A. Lankton, A. Tannenbaum . Localizing Region Based Active Contours. IEEE Trans.Image Process,Vol. 17,Issue 11 (2008) 2029-2039

[12] J.A. Yezzi, A. Tsai, A. Willsky. A Fully Global Approach to Image Segmentation via Coupled Curve Evolution Equations. J. Vis. Comm. Image Rep ,Vol.13,Issue 1 (2001) 195-216

[13] W. Aitfares, A. Herbulot, M. Devy, E.H. Bouyakhf, F. Regragui, F. A Novel Region-Based Active Contour Approach Relying on Local and Global Information. In proc. of IEEE International Conference on Image Processing (ICIP), Bruxelles, Belgique (2011) 1049-1052

[14] [12] W. Aitfares, E.H. Bouyakhf, F. Regragui, A. Herbulot, M. Devy. A Robust Region-based Active Contour for Object Segmentation in Heterogeneous Case. Pattern Recognition and Image Analysis journal,Vol. 24, Issue 1 (2014) 24-35

[15] W. Aitfares, A. Herbulot, M. Devy, E.H. Bouyakhf, F. Regragui.Une Nouvelle Approche des Contours Actifs Basés Région Utilisant une Information Locale et Globale. In proceedings of 23éme Colloque GRETSI sur le traitement du signal et des images, Bordeaux, France, Septembre (2011)

[16] C. Harris, M. Stephens. A Combined Corner and Edge Detector. 4 TH + Alvey vision conf. Manchester (1988) 189+.

[17] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems Man and Cybernetics (1979) 62-66

[18] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong. ‘A new method for gray-level picture thresholding using the entropy of the histogram. Graph. Models Image Process. 29, 273–285 ~1985.

[19] J. Bernsen. Dynamic thresholding of gray level images. ICPR’86: Proc. Intl. Conf. Patt. Recog., pp. 1251–1255 ~1986.

[20] C.K. Chow, T. Kaneko. Automatic detection of the left ventricle from cineangiograms. Comput. Biomed. Res., 1972, 5, pp. 388–410.

[21] L. Eikvil, T. Taxt, K.A. Moen. Fast adaptive method for binarization of document images. Proc. ICDAR, France, 1991, pp. 435–443.

[22] K.V. Mardia, T.J. Hainsworth. A spatial thresholding method for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell., 1988, 10, (8), pp. 919–927.

[23] W. Niblack. An introduction to digital image processing. Prentice- Hall, Englewood Cliffs, NJ, 1986, pp. 115–116.

[24] T. Taxt, P.J, Flynn, A.K. Jain. Segmentation of document images. IEEE Trans. Pattern Anal. Mach. Intell., 1989, 11, (12), pp. 1322–1329.

[25] S.D. Yanowitz, A.M. Bruckstein. A new method for image segmentation.Comput. Vis. Graph. Image Process., 1989, 46, (1), pp. 82–95.

[26] J. Sauvola and M. Pietaksinen.Adaptive document image binarization.*Pattern Recogn.* **33**, 225–236 (2000).

[27] H. Moravec. Obstacle Avoidance and Navigation in The Real World by a Seeing Robot Rover. Technical Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute (1980).

[28] Harris, C., Stephens, M.: A Combined Corner and Edge Detector. 4 TH + Alvey vision conf. Manchester (1988) 189+.

[29] J. Song, R. L. Stevenson, E. J. Delp, 1989. The Use of Mathematical Morphology in Image Enhancement", Proceedings of the 32nd Midwest Symposium on Circuits and Systems, Urbana-Champaign, IL, 67-70

[30] The Berkley Segmentation Database and Benchmark. [online] http://www.eecs.berkley.edu/-Research/Projects/CS/vision/bsds/.C

[31] Rother, V Kolmogorov and A Blake. GrabCut: interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics (TOG), vol.23, pp. 309-314, 2004.

Biographies

Nada El aryby received the B.Sc and the M.Sc degrees from Mohammed V University, Rabat, Morocco, respectively, in Physics in 2010, and Informatics,Signals and Telecommunications in 2012. Ph.D student at Mohammed V University, Morocco in the LIMIARF Laboratory under the supervision of El Houssine Bouyakhf. Her research interests include Image segmentation, Object detection and tracking.

Wassima Aitfares received the B.S and the M.Sc degrees from Mohammed V University, Rabat, Morocco, respectively, in Mathematics in 2005, and Informatics, Telecommunications and Multimedia in 2007. She received her Ph.D. degree from Mohammed V University, Morocco with collaboration with Paul Sabatier University, Toulouse, France in 2013. Her research interests include Segmentation, Detection, Object Tracking and the Analytic Videos.

Fakhita Regragui received the Master degree in Electronics from the Faculty of Sciences at Mohammed V University –Agdal, Rabat, Morocco in 1977 and the PhD degree in Signal processing and Communications from Rutgers University, New Brunswick, New Jersey, USA in 1990. Worked, from 1977 to 1983, as Assistant professor at the Faculty of Sciences teaching courses related to the Electronics field. Since 1990, she worked as Full Professor at the same Faculty teaching Electronics and Digital Signal Processing courses respectively at undergraduate and graduate levels. Her research interests include applications of signal and image processing in the biomedical field. In signal processing, she is currently working on filtering and classification using different techniques (adaptive filtering, wavelet analysis, neural networks,..) and their applications to neuroscience. In the field of image processing, she is currently working on different tools such as contour detection and texture analysis for different image processing applications (mammograms interpretation, estimation of fish growth parameters, motion tracking applied to robotics).